The Thesis committee for Nicholas Robert Griffin

Certifies that this is the approved version of the following thesis:

The Effects of Depressive Symptoms on Memory Distortion for Orthographic Associates: A Behavioral and EEG Investigation

APPROVED BY

SUPERVISING COMMITTEE:

David M. Schnyer, Supervisor

Jessica A. Church-Lang

The Effects of Depressive Symptoms on Memory Distortion for Orthographic Associates: A Behavioral and EEG Investigation

by

Nicholas Robert Griffin

Thesis

Presented to the Faculty of the Graduate School

of the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

August 2017

The Effects of Depressive Symptoms on Memory Distortion for Orthographic Associates: A Behavioral and EEG Investigation

by

Nicholas Robert Griffin, M.A.

The University of Texas at Austin, 2017

SUPERVISOR: David M. Schnyer

While everyone is prone to memory errors, individuals with cognitive biases like negative attention bias may differ in their susceptibility to memory distortion. Negative attention bias, which is frequently comorbid with depressive symptoms, has been associated with increased false memory for negative stimuli. Across two experiments, we used lists of orthographically associated words to analyze responses to novel emotional stimuli between individuals with depressive symptoms and healthy controls. In Experiment 1, participants encoded neutral words that were orthographically associated with neutral, negative, or positive critical lures. Then, they completed a recognition memory test on the words shown during encoding (old items) and the critical lures (novel items). We did not find differences in false alarms to novel stimuli between groups. However, we did find significant differences in response times to correct rejections of novel stimuli and greater false alarm confidence in the depressive symptom group. In Experiment 2, we used electroencephalography to further investigate the mechanisms through which each group may complete this task. Using only lists associated with negative and neutral critical lures, the depressive symptom group showed greater hit and false alarm rates, and quicker overall response times than controls. The EEG results suggested that the behavioral differences might arise from separate memory retrieval strategies, with the depressive group utilizing an earlier, familiarity-based retrieval strategy and the control group utilizing a later, recollection-based retrieval strategy. These experiments are an important first step in understanding mechanisms underlying the relationship between depressive symptoms and emotional stimuli on memory distortion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Contrary to common perceptions about the validity of our memories, they are instead highly prone to distortion. Over time, memories for events become less precise, requiring individuals to fill in increasingly large gaps in memory as time elapsed since encoding increases[1,2]. Further, subsequent events can influence memory for an original event and impact how it will be retrieved in the future[3–5]. Elucidating how memories become distorted is critical not only to understand the susceptibility to incorporate misinformation into memory, but also to reduce retrieval of false information. A range of literature has examined methods through which memory can become distorted, including suggestion of misinformation[3,6–8] and semantic priming in word learning tasks[9,10]. Although many studies show strong memory distortion effects for healthy samples, additional factors including biased attention may further affect the extent to which memory may become distorted following encoding, making those with mental disorders particularly vulnerable.

One such factor affecting memory distortion is negative attention bias, or the tendency to preferentially attend to negative stimuli[11,12]. Negative attention bias, which may be measured through extended response times to negative stimuli, is highly comorbid with depressive symptoms or with clinical diagnoses of Major Depressive Disorder[13–16]. Understanding the effects of negative attention bias is critical for informing interpretations of existing false memory research, which commonly uses emotional stimuli to elicit endorsements of misinformation[3,17]. Further, with depression ranking as one of the most commonly diagnosed psychological disorders[18], expansion of the negative attention bias research will allow for a greater understanding of the cognitive biases within the considerable subset of the population who experience depressive symptomatology.

Several studies have used emotional stimuli to examine the effect of negative attention bias on subsequent memory retrieval. Such experiments designed for testing memory retrieval accuracy for emotional stimuli have shown that, due to increased attention to negative items, correct recall for negative stimuli is greater than neutral[19] and positive stimuli[20]. Additional studies that have specifically examined the effects of emotion using the Deese-Roediger-McDermott (DRM) paradigm for memory distortion have shown an opposite effect[9]. In DRM experiments, participants learn lists of words that are semantically associated with a theme word called a critical lure. For example, a list of words like “bed,” “rest,” and “dream,” would all be related to the critical lure “sleep.” Subsequently, participants must recall previously encoded words, which often leads to falsely recalling (non-encoded) critical lures. One recent DRM study that used lists associated with positive, negative, and neutral critical lures found that participants diagnosed with Major Depressive Disorder were more likely to falsely recall negative critical lures than control participants[21]. A similar study that tested participants on list words and critical lures via a recognition memory test found greater false alarms to negative critical lures, and suggested that increased semantic similarity of negative information drove their results[22]. While the results from these studies show how negative biases may influence memory retrieval outcomes, additional research is needed for a fuller understanding of the effects of emotion on word learning memory distortion paradigms.

To better examine the effects of emotion on memory recall, we used an adaptation of the DRM paradigm that aimed to reduce emotion-based priming for non-neutral test stimuli. While prior studies have repeatedly shown strong memory distortion effects from the DRM[10], emotion manipulations are inherently confounded. In standard DRM lists, all list words are of the same valence as their associated critical lure. Therefore, encoding the words from one DRM list would prime subjects for the presentation of novel items of similar valence during test. Such priming may artificially inflate the false alarm rate for critical lures during test, and particularly for negative critical lures. Thus, to maximize our control of emotion, we instead created lists of orthographically associated words. This adaptation allows for the critical lures and list words to have different valence; emotional words can be isolated to critical lures, while the list words can be emotionally neutral.

Two prior studies have utilized lists of orthographically associated words to test memory for emotional stimuli. In the first study, participants encoded words from 12 lists of neutral orthographic associates that were related to either neutral or negative critical lures[23]. Like DRM tasks, the subsequent tests included both the list words and the critical lures for each encoded word list. The researchers found that false alarm rates were greater for the neutral than the negative critical lures. Further, they found that false alarms were reduced when emotional items were included as encoding stimuli alongside the existing neutral encoding items[23]. Finally, they found that although false alarm rates were greater for neutral than negative critical lures, confidence was greater for the false endorsement of negative than neutral lures. A similar study with older adults replicated the relatively low false alarm rate for negative critical lures and found significantly more false alarms to positive critical lures than both negative and neutral lures[24]. Although reduced false alarm rate has been shown for negative items in an orthographic associates task, the effects of depressive symptoms and negative attention bias in this design has not yet been explored.

In the present study, we investigated the effects of negative attention bias on memory distortion for emotional stimuli using lists of orthographically associated words. Critically, we aimed to examine responses to emotional stimuli by individuals with depressive symptoms without semantic priming, and instead preserved the novelty of emotional items during the subsequent memory test. To test effects of high or low negative attention bias, we conducted two experiments with two groups of participants: one group with elevated depressive symptoms (i.e. greater negative attention bias) and one group of healthy controls (i.e. lower negative attention bias). In our first experiment, we participants completed a behavioral recognition memory task using orthographically associated word lists to elucidate memory retrieval differences between the groups. We hypothesized that the overall pattern of false alarms would mirror previous findings, such that false alarms from the control group would be greater for neutral than negative critical lures. Conversely, we hypothesized that due to negative attention bias, the depressive symptom group would show greater false alarms for negative than neutral critical lures. In our second experiment, we conducted a similar experiment while collecting electroencephalography (EEG) data to determine potential differences in cortical activation patterns during memory retrieval.

Experiment 1.

Participants. We recruited 64 students (47 females, 17 males; mean age = 19.9 years, SD = 1.9) from the Psychology 301 subject pool at the University of Texas at Austin. Participants were given class credit for their participation.

Materials. Participants encoded individual words from orthographically associated word lists as they were presented on a monitor. Each list also contained a theme word, or critical lure, which served as the reference for all associate words in a given list, but which was not shown during encoding. While the valence of the critical lures varied, the valence for all associate list words was neutral, regardless of lure valence. For example, a list with the positive critical lure "king" contained orthographically associated words such as "ping," "ring," and "wing." During encoding, participants only viewed list words; list words and critical lures were both presented during subsequent testing. Word lists and critical lures were derived from previous studies using the orthographic associates DRM paradigm[23,24] to study emotion. Additional lists were generated by selecting critical lures from words that were previously normed for affect, critically including ratings for valence[25].

Throughout the task, participants viewed words from 24 lists; eight lists of words each related to neutral, negative, or positive critical lures. In addition to the list words, participants viewed the associated critical lures during test. Therefore, participants were tested on 240 total list words and 24 critical lures (see Appendix A for full Exp. 1 word lists).

Encoding and testing were divided into two blocks to reduce testing fatigue; each block consisted of list word encoding, a distractor task to reduce information rehearsal, and a test of list words and critical lures. In one block, participants encoded list words associated with neutral and negative critical lures, and in the other block, participants encoded list words associated with neutral and positive critical lures. The presentation order of the blocks was counterbalanced across participants.

To measure depressive symptoms, participants completed the Center for Epidemiologic Studies – Depression scale (CESD)[26]. The CESD consists of 20 items that are designed to assess depressive symptoms over the past week. Total scores range from 0 (least depressed) to 60 (most depressed). Although prior studies have used a score of 16 as the threshold for elevated depressive symptoms[27], we selected a scoring threshold of 20 to increase our confidence that our sample was truly representative of individuals with depressive symptoms. Based on the CESD scores from the present experiment, 32 participants scored into the depressive symptom group (CESD mean = 28.8, SD = 7.1), and 32 participants scored into the healthy control group (CESD mean = 9.34, SD = 3.3).

For each analysis of recognition memory results (hits and false alarms), we conducted a 3 X 2 ANOVA with list valence (positive critical lure, neutral critical lure, and negative critical lure) and participant group (healthy control [HC] and depressive symptom [DP]) as factors. For analysis of response times, we conducted a 3 X 2 X 2 ANOVA with list valence, participant group, and accuracy (hits or misses for encoded items, correct rejections of false alarms for novel items) as factors. For analyses with a significant interaction between factors, we conducted additional lower-order ANOVAs or t-tests to post-hoc elucidate the differences between factors. For significant main effects of list valence, we conducted t-tests between the three list categories to elucidate differences between categories. To eliminate outlier responses, trials with response times less than 300ms and response times greater than two standard deviations above the mean response time for each participant were excluded from the analyses. Finally, we conducted two 3 X 2 X 2 ANOVAs for each analysis of confidence ratings at test, with list valence, group, accuracy as factors. One model was conducted for responses to list words (encoded words) and the other model for responses to critical lures (novel test items). See Appendix B for full summary statistics.

Procedure.

Pretest. Participants first completed a pretest to gain exposure to the format of the experimental tasks. The pretest consisted of randomized, individually presented words from one of two unique word lists. Both lists contained ten words that were orthographically associated with neutral critical lures. During encoding, each word was presented for 2500ms, with a 500ms interstimulus interval. A fixation cross was presented at the center of the screen during the interstimulus interval. Immediately following encoding, participants completed a recognition memory test on the previously encoded list words and the critical lures for both lists. Test items were presented individually, and participants had to endorse each item as either old (previously encoded) or new (novel item). Participants were given as much time as they needed to complete the test.

Encoding. Participants completed two encoding blocks in which words from sets of eight lists of orthographic associates were randomly and individually presented. One encoding block presented four lists associated with positive critical lures and four lists associated with neutral critical lures. The other block presented four lists associated with negative critical lures and four lists associated with additional neutral critical lures. The presentation order of the encoding blocks was counterbalanced across participants. Words from each list in the given encoding block were randomized and individually presented for 2500ms with a 500ms interstimulus interval. A fixation cross was presented at the center of the screen during the interstimulus interval. Each list presented at encoding contained ten words, for a total of 80 words presented during each encoding block. Across both encoding blocks, participants encoded a total of 160 words. Participants viewed each encoding stimulus once before advancing to the next task.

Distractor task. To reduce information rehearsal between each of the encoding and test blocks, participants completed two distractor tasks. Between the first encoding and test blocks, participants completed Operation SPAN[28]. During this task, participants viewed a series of arithmetic problems and were asked to judge each problem as true or false. Then, a single letter was shown on the screen that participants were asked to remember. After each block containing three to seven trials of arithmetic and letter memorization, participants had to recall each letter from the most recent series in the order they were presented. Between the second encoding and test blocks, participants completed a bubble shooter game[29] which required them to group same-colored bubbles together to earn points. Duration of the bubble shooter game was 5-10 minutes.

Test. Following each distractor task, participants completed an old/new recognition memory test for the items shown during the most recent encoding block. Additionally, the critical lures associated with the items shown during encoding were included in the recognition test. Finally, words from four novel lists of orthographic associates were included in the test. The valence of the critical lures of novel lists matched the valence of the non-neutral lists presented at encoding. For example, if the encoding block included lists of words associated with negative and neutral critical lures, then the novel lists of words were all associated with negative critical lures. In each testing block, participants endorsed 132 words as either old or new (80 from encoding, 8 critical lures associated with the encoded words, 40 novel list words, and 4 critical lures associated with the novel list words). Across both blocks, participants were tested on 264 total words.

Following the old/new endorsement of each test item, participants rated their confidence in their old/new response. Confidence was rated on a 5-point Likert scale from 1 (not at all confident) to 5 (very confident). Participants were given as much time as they needed to complete both the old/new endorsement and the confidence rating for each item.

Finally, participants completed the CESD and a demographic questionnaire. The CESD was completed at the end of the session to reduce the potential introduction of negative bias following the completion of a measure of depressive symptoms.

Results.

We first tested effects of list valence and group on correct recognition (hits) for items encoded before the test blocks. We did not find significant main effects of list valence, $F(2,186) = .53$, $p = .59$, $\mu_p^2 = .01$, or group, $F(1,186) = 1.70$, $p = .19$, $\mu_p^2 = .03$. The interaction between the two factors, however, was trending toward significance, $F(2,186) = 2.43$, $p = .09$, $\mu_p^2 = .03$. The pattern of the interaction was such that the HC group showed slightly fewer correct hits for neutral than negative items, while the DP group showed slightly *greater* hits for neutral than negative items (Figure 1).

Next, we tested the effects of list valence and group on false to critical lures, which had not been shown before the testing block. We found a significant main effect of list valence, $F(2,186) = 36.71$, $p < .001$, $\mu_p^2 = .28$ (Figure 1; Table 1). Then we conducted a series of t-tests to determine the differences in false alarm rates between each list valence. We found that false alarm rates to neutral critical lures were significantly greater than false alarm rates to both negative, $t = 7.74$, $p < .001$, 95% CI = [.23, .38], and positive critical lures, $t = 6.74$, $p < .001$, 95% CI = [.19, .36]. False alarm rates between positive and negative critical lures were not significantly different, $t = .82$, $p = .41$, 95% CI = [-.04, .11]. The main effect of group, $F(1,186) = 2.22$, $p = .14$, $\mu_p^2 = .01$, and the interaction between the factors, $F(2,186) = .56$, $p = .57$, $\mu_p^2 = .006$, did not reach statistical significance.

In the analysis of response times to previously encoded items (hit or miss responses), we found a significant main effect of accuracy, with quicker response times for correct than incorrect responses, $F(1,370) = 20.18$, $p < .001$, $\mu_p^2 = .05$. We also found that the main effect of group was trending toward significance, with slightly quicker overall response times in the HC group than the DP group, $F(1, 370) = 2.93$, $p = .09$, $\mu_p^2 = .008$. The main effect of list valence, however, was not significant, $F(2, 370) = .32$, $p = .73$, $\mu_p^2 = .002$. In addition, none of the second-order interaction terms between the factors returned significant results, F's < 1.4, p's > .25. Finally, the three-way interaction term including all three factors was not statistically significant, $F(2, 370) = .04$, $p = .96$, $\mu_p^2 < .001$.

For the analyses of response times to critical lures, we did not find significant main effects of group, $F(1,348) = 2.25$, $p = .13$, $\mu_p^2 = .006$, list valence, $F(2, 348) = .42$, $p = .66$, $\mu_p^2 = .002$, or accuracy, $F(1,348) = .79$, $p = .38$, $\mu_p^2 = .002$. Further, we did not find a significant interaction between group and list valence, $F(2, 348) = .25$, $p = .78$, $\mu_p^2 = .001$, or between group and accuracy, $F(1, 348) = .06$, $p = .81$, $\mu_p^2 < .001$. However, we did find a significant interaction between list valence and accuracy, $F(2, 348) = 4.17$, $p = .02$, $\mu_p^2 = .02$ (Table 1). Additional post-hoc examination of the list valence showed no significant differences in response time between valence of false alarm responses, t's < 1.3, p's > .2. However, differences in response times for negative and neutral correct rejections were trending towards significance, $t = 1.65$, $p = .1$, 95% CI = [-.03, .32], with quicker response times for negative items. Furthermore, we found significantly quicker response times for positive relative to neutral correct rejections, $t = 3.31$, $p = .001$, 95% CI = [.11, .440], and for positive relative to negative correct rejections $t = 2.21$, $p = .03$, 95% CI = [.01, .25] (Figure 2). Finally, we found that the three-way ANOVA including group, list valence, and accuracy was trending toward significance, $F(2, 348) = 2.33$, $p = .10$, $\mu_p^2 = .01$. This effect was mainly driven by the DP group, which showed quicker response times for neutral than emotional false alarms and slower response times for neutral than emotional correct rejections. However, because of the small effect size of the interaction, this finding should be interpreted cautiously.

Our final set of analyses focused on the confidence ratings following each test trial (Table 1). In the analysis of confidence ratings to list words, we found significantly greater confidence for correct (hit) responses than incorrect (miss) responses, $F(1, 370) = 119.53$, $p < .001$, $\mu_p^2 = .24$. We also found a significant effect of list valence, $F(2, 370) = 3.38$, $p = .03$, $\mu_P^2 = .02$, with greater confidence for neutral than negative, $t = 1.86$, $p = .06$, 95% CI = [-.01, .36], and positive items, $t = 1.93$, $p = .054$, 95% CI = [-.003, .38], narrowly missing significance. Confidence was not significantly different between negative and positive items, $t = -.12$, $p = .91$, 95% CI = [-.18, .16]. Interestingly, we also found significantly *lower* confidence in the DP group than the HC group, $F(1, 370) = 5.28$, $p = .02$, $\mu_p^2 = .01$. The interaction between list valence and accuracy was also significant, $F(2, 370) = 5.59$, $p = .004$, $\mu_p^2 = .03$, driven by high confidence ratings for hits to neutral relative to non-neutral items and confidence for miss responses regardless of valence. We did not find significant interactions of list valence and group, $F(2, 370) = .02$, $p = .98$, $\mu_p^2 < .001$, or group and accuracy, $F(1, 370) = 1.82$, $p = .18$, $\mu_p^2 = .005$. The interaction term including all three main factors also did not reach significance, $F(2, 370) = .04$, $p = .96$, $\mu_p^2 < .001$.

In our analysis of confidence ratings for critical lures, we found a significant effect of list valence, $F(2, 348) = 4.01$, $p = .02$, $\mu_p^2 =.02$, such that confidence was greater for responses to negative lures than neutral or positive lures (Table 1). The main effects of group, $F(1, 348) = .25$, $p = .62$, $\mu_p^2 < .001$, and accuracy, $F(1, 348) = .59$, $p = .44$, $\mu_p^2 = .002$, were not significant. We did find a significant interaction between group and accuracy, $F(1, 348) = 8.15$, $p = .005$, $\mu_p^2 = .02$, with the DP group showing greater confidence in false alarms than correct rejections, and the HC group showing the opposite response pattern. In addition, we found a significant interaction of list valence and accuracy, $F(2, 348) =10.84$, $p < .001$, $\mu_p^2 = .06$. For this interaction, confidence was greater for neutral than non-neutral false alarms, while confidence was greater for negative correct rejections than neutral and positive correct rejections. Neither the interaction between list valence and group, $F(2, 348) = .78$, $p = .46$, $\mu_p^2 = .004$, nor the interaction between all three factors, $F(2, 348) = .68$, $p = .51$, $\mu_p^2 = .004$, reached significance.

Table 1. Experiment 1 – Significant Behavioral Results

Exp. 1 - Significant Behavioral Results		Healthy Control Group		Depressive Group	
	Valence	Mean	SD	Mean	SD
Critical Lure False Alarm Rate	Positive	0.25	0.20	0.34	0.23
	Neutral	0.54	0.25	0.60	0.24
	Negative	0.26	0.21	0.27	0.21
False Alarm (FA) Response Time	Positive	1.49	0.59	1.80	0.86
	Neutral	1.55	0.46	1.45	0.52
	Negative	1.53	0.55	1.68	1.05
Correct Rejection (CR) Response Time	Positive	1.49	0.29	1.51	0.35
	Neutral	1.66	0.42	1.89	0.69
	Negative	1.62	0.39	1.64	0.34
Hit Confidence Rating	Positive	3.84	0.66	3.77	0.46
	Neutral	4.20	0.64	4.18	0.45
	Negative	3.80	0.64	3.71	0.55
Miss Confidence Rating	Positive	3.30	0.74	3.04	0.70
	Neutral	3.28	0.68	3.04	0.75
	Negative	3.35	0.59	3.14	0.79
FA Confidence Rating	Positive	3.32	1.09	3.48	0.96
	Neutral	3.57	0.87	3.81	0.90
	Negative	3.25	1.09	3.55	0.96
CR Confidence Rating	Positive	3.61	0.84	3.32	0.80
	Neutral	3.54	0.78	2.94	1.02
	Negative	4.04	0.79	4.01	0.70

Figure 1. Experiment 1: Mean Hit and False Alarm Rates

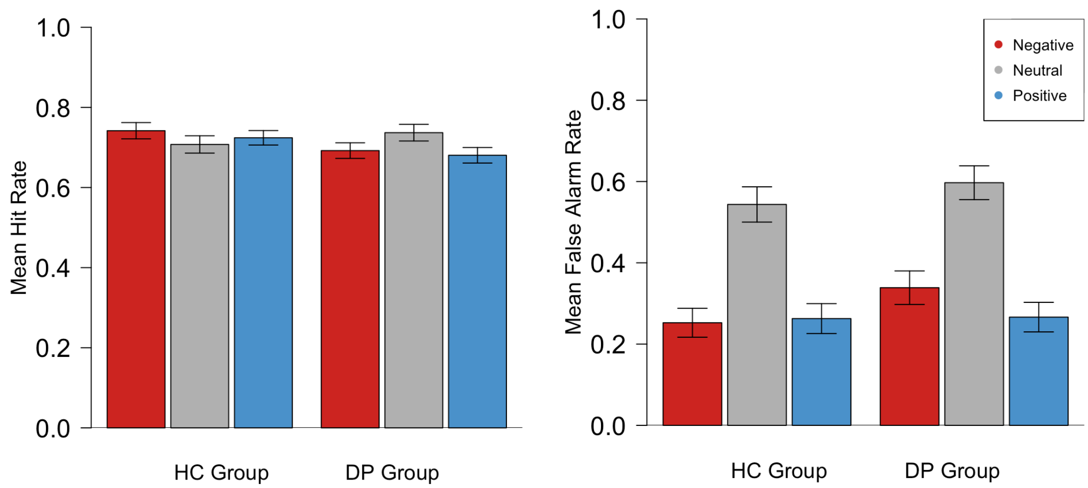

Figure 1. On the left: mean hit rates (±1SE) in Experiment 1 to list words by group and valence of the associated critical lure. On the right: mean false alarm rates (±1SE) to critical lures by group and valence.

Figure 2. Experiment 1: Mean Correct Rejection Response Times

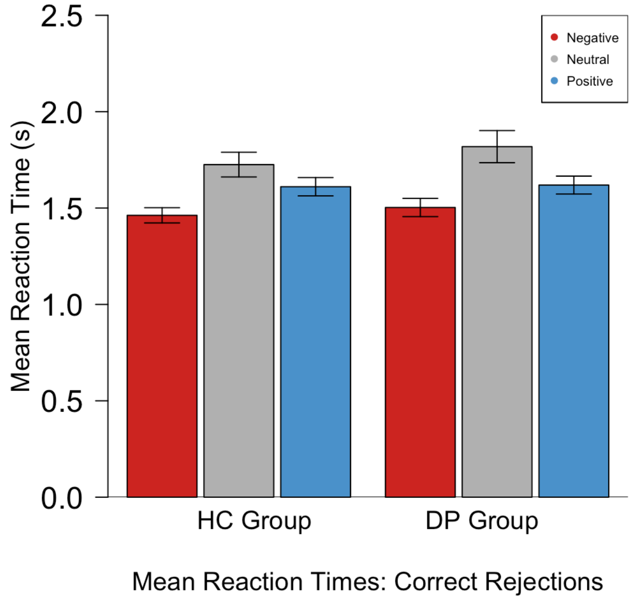

Figure 2. Mean response times (±1SE) in Experiment 1 for correct rejections of critical lures by group and valence.

Interim Discussion

Contrary to the hypothesis, we did not find significant differences between HC and DP groups for the emotional critical lures. Instead, both groups performed similarly in their memory performance for the lures: greater false alarm rates for neutral than emotional lures, and similar false alarm rates for lures with positive and negative valence. However, the data does provide evidence that the groups processed some of the test items differently. For example, we found a marginally significant interaction between the two groups and valence in recognition for previously encoded items. With the depressive symptoms group showing a numerically greater hit rate to neutral items than the control group, the former group may have been somewhat more likely to endorse test items as old (previously encoded). Furthermore, we found differing response times in both groups between list valences when correctly rejecting critical lures. With both groups rejecting positive lures more quickly than lures of other valences, the positive items may have more arousing, allowing for quicker rejection during test. Perhaps more informative were the results of the confidence ratings, with the depressive group showing greater confidence for false alarms than correct rejections, and the control group showing the opposite pattern. This interaction of confidence scores for critical lures was not mirrored in recognition performance, suggesting that the relationship between confidence and recognition memory requires further examination. The complexity of our behavioral findings and the number of marginal results prompted us to further investigate design components that may have affected our results, and the mechanisms underlying the behavior from each group.

One of the most interesting findings from the first experiment was the lack of valence effects in recognition and response time, specifically for the depressive symptom group. The lack of significant valence effects may have been rooted in one or more limitations. First, the low number (eight) of emotional critical lures during each test block means that the differences in false alarms between participants is not very large; a greater number of critical lures would provide more statistical power. Secondly, the positive and negative critical lures were presented in separate encoding and testing blocks, and they were the only emotional stimuli within each block. Thus, responses to non-neutral items may have been more salient during test regardless of valence. To address these issues, we aimed to create more critical lures, and thus more lists, in the subsequent experiment. We also aimed to include each valence of stimuli in each experimental block. However, according to Smith et al. (2006), behavioral outcomes resulting from negative attention bias may be reduced to the point of non-significance when positive stimuli are presented alongside negative stimuli[30]. Therefore, we removed word lists associated with positive critical lures to reduce potential attenuation of valence-related memory effects.

The significant results from the confidence ratings in Experiment 1 inspired an additional question: whether differences in cortical activation exist between groups as they complete the recognition test. Prior EEG research has shown that confidence in a behavioral response positively correlates with amplitude of event-related potentials associated with that response[31,32]. For example, greater confidence on a memory test would elicit more extreme amplitudes of the components associated with memory retrieval, like the N400 and late positive component (LPC). Certain confidence rating results from Experiment 1 suggest that the groups may have different patterns of activation, like the greater confidence reported by the control group to previously encoded items than the depressive symptom group. Additionally, the depressive symptom group was more confident in false alarm responses than correct rejections of critical lures, which the control group had *less* confidence for false alarms. These findings suggest that the two groups may have different cortical responses to falsely endorsing novel stimuli. Combined with the elimination of positive stimuli, potential differences in cortical activation may be more easily assessed and understood in the context of negative attention bias.

In the second experiment, we used EEG recording to identify the patterns of cortical activation for healthy controls and individuals with depressive symptoms using the orthographic associates task with only neutral and negative items. We also made several changes to the experiment design to optimize the task for EEG data collection. First, we created additional lists related to neutral and negative critical lures. Each list was shortened from ten to eight words long to increase the ratio of critical lures to list words during testing. To compensate for the increased time required to complete the experiment due to the greater number of stimuli, we removed the distractor tasks between encoding and testing blocks. Finally, to ensure that participants were attending to the stimuli presented during encoding, we added a shallow encoding (i.e. low-level processing)[33] task for participants to complete as stimuli were presented.

For the behavioral outcomes, we expected to find a stronger effect of valence between the groups. Specifically, we hypothesized that the depressive symptom group would have fewer false alarms than the control group for negative critical lures. We predicted that this effect would be driven by increased reactivity to negative stimuli in the depressive symptom group, allowing them to reject novel negative items more effectively than controls. For the EEG analysis, we hypothesized that the depressive symptoms group would have greater cortical activation for negative stimuli in early components associated with memory retrieval. Prior research has shown heightened reactivity to negative stimuli very shortly after stimulus onset (within 200ms)[34]. While this more extreme amplitude has been noted in non-clinical populations, individuals with negative attention bias, like those with depressive symptoms, may show a more acute response to negative stimuli. Due to the response time results of Experiment 1, we also hypothesized that the control group would show stronger cortical activation for later post-stimulus retrieval components than the depressive symptom group. According to the dual process theory of memory retrieval[35–37], early components of memory retrieval are associated with familiarity processes. In contrast, later retrieval-related activation is associated with more contextually rich recollection processes. Therefore, we hypothesized that participants with depressive symptoms may show more familiarity-related cortical activation during recognition memory testing, while the healthy control group would show more recollection-related activation.

Experiment 2.

Participants. We recruited 73 students from the Psychology 301 subject pool at the University of Texas at Austin. Participants were given class credit for their participation. From the 73 participants, 27 were excluded from the analyses because of test accuracy at or below chance (n=11), EEG data that was too noisy to analyze (n=6), and unstable CESD scores, which caused inconsistent group categorization (n=10). Thus, 46 participants were included in the final analyses (31 females, 16 males; mean age = 19.7, SD = 1.5).

Materials and Equipment. Participants encoded individual words from lists of words that were orthographically associated with either a neutral or a negative critical lure as they were presented on a monitor. To improve the EEG signal to noise ratio (i.e. increase the number of critical lures relative to list words), participants encoded and were tested on words from 40 word lists. Each list consisted of 8 orthographic associates, and 20 lists were associated with neutral critical lures while the other 20 lists were associated with negative critical lures (see Appendix C for full Exp. 2 word lists). Participants indicated whether each word had an even or odd number of vowels as they were presented during encoding. The stimuli were divided into four encoding blocks, with each block consisting of five word lists associated with neutral critical lures and five word lists associated with negative critical lures. Within each block, the list words were randomized and individually presented for 2500ms with 500ms interstimulus intervals. A test block immediately followed each encoding block. During each test block, participants completed an old/new recognition memory test for the list words and the associated critical lures from the most recent encoding block. Presentation order of the encoding and testing blocks was randomized for each participant.

The CESD was again used to measure depressive symptoms and determine group placement. Consistent with existing literature, we used a score of 16 as the threshold for placement into the high depressive symptom group. All participants completed the CESD at the end of their experimental session, after EEG data collection and memory testing. Based on the CESD scores of the 46 participants whose data were included in the final analyses, 25 scored into the depressive symptom group (mean = 22.2, SD = 6.6) and 21 scored into the healthy control group (mean = 8.1, SD = 3.8).

Participants also completed the Karolinska Sleepiness Scale[38] (KSS) at the beginning of each encoding block which allowed us to track fatigue over the duration of each experiment session. The KSS is an integer scale ranging from 1 (not at all tired) to 9 (extremely tired).

For the analyses of behavioral outcomes, we first tested for any differences in the responses to list words during encoding. To do this, we conducted a 2 X 2 ANOVA with list valence (list words associated with negative or neutral critical lures) and participant group (healthy control [HC] and depressive symptom [DP]) as factors. Thus, we could determine if behavior during encoding contributed to test outcomes. For recognition memory outcomes, we conducted 2 X 2 ANOVAs with list valence (neutral critical lure and negative critical lure) and participant group as factors. Because of the unequal group sizes, all ANOVAs were completed using Type III sum of squares. For the analyses of response times, we conducted 2 X 2 X 2

ANOVAs with accuracy (hits and misses, or correct rejections and false alarms), list valence, and participant group as factors. In all analyses, trials with response times less than 300ms and response times greater than two standard deviations above the mean response time for each participant were excluded. See Appendix B for full summary statistics.

Continuous EEG data was collected throughout encoding and testing with a 64-channel BioSemi active electrode cap, plus one electrode placed on each mastoid. Locations of the electrode channels were based on the extended 10/20 location system. In addition, four electrodes were placed on the face (one below each eye and one lateral to the canthi of each eye) to measure bilateral vertical and horizontal eye movements. All channels were amplified with a BioSemi Active II amplifier, and all channel impedances were kept below the BioSemi recommended maximum threshold of ± 40 μV. As EEG data was recorded, participants sat in a sound-insulated, low-light environment and completed the experimental tasks as shown on a monitor in front of their seat. Responses were recorded through a keyboard that was placed in the lap of each participant for the duration of the experimental tasks.

Preprocessing of EEG data was completed with BrainVision Analyzer. Data for 39 participants was band pass filtered between 0.1 and 30 Hz, and data for 7 participants was band pass filtered between 1 and 30 Hz to eliminate a relatively greater level of electrical interference. No more than three EEG channels were interpolated for each participant; data that required more than 3 channels to be interpolated were excluded from the analyses. Each channel was then re-referenced to the linked mastoids. By averaging the horizontal and vertical EOG channels, respectively, two offline bipolar ocular channels were created to remove ocular motion artifacts. Then, the data were segmented into epochs of 1700ms (-200 to 1500ms about stimulus onset of each trial). For each epoch of the EEG data, DC drift was corrected and a baseline correction was completed in reference to the 200ms interval before stimulus onset.

The EEG data were initially analyzed using pointwise non-parametric randomized permutation t-tests, with multiple comparisons corrections across time and electrode location[39,40]. This type of analysis allows for the elucidation of clusters of significant activation differences between two groups, including small but significant clusters of activation. The permutation technique is more conservative than traditional EEG testing methods, and with an analysis of the entire scalp it does not require an a priori region of interest to be specified.

The pointwise permutation method requires three steps to determine the size and spatiotemporal location significant clusters. First, a statistical significance threshold was computed for each electrode location and time point. With a total of 64 active electrodes recording 1.7 sec (1700ms) epochs at a 256 Hz sampling rate, we computed 27,840 independent thresholds. The thresholds were determined using an estimated t-distribution that computed t-statistics from 20,000 random permutations of the data under the null hypothesis. The t-statistics comprising the null distribution were generated after permuting the data of a randomized subset of participants in each group. Thus, a three-dimensional matrix was created including two spatial dimensions, and a third dimension of time. From that matrix, the locations of significant t-values are used to determine locations of clusters in which activation exceeds significance thresholds.

The second step of the analysis corrects for the inflated type-I error from the large number of independent t-tests. To do this, the locations of significant t-values were used to determine the locations of contiguous clusters of significance within the three-dimensional matrix. A second round of 20,000 permutations was completed to determine the null distribution of clusters that collectively exceed significance. We computed the exceedance mass for each cluster, or "the integral of the statistic image above the primary threshold within the suprathreshold cluster"[40]. The value of the largest exceedance mass in each permutation was used to create a distribution of maximal exceedance values under the null hypothesis.

Finally, this second distribution of suprathreshold cluster is tested against the t-statistic clusters from the non-permuted dataset. This test was conducted with a Holm-Bonferroni step-down correction of the null distribution's p=.05 criterion exceedance mass. True clusters with exceedance mass greater than p=.05 will be considered significant in a two-tailed analysis. Together, this technique allowed us to identify the locations and times following stimulus onset that statistically significant differences in cortical activation occurred. For the present study, we were interested in significant clusters arising in the medial frontal and left parietal cortex in accordance patterns of cortical activation for attention processing and memory retrieval.

Once the clusters were determined for each of the within-subjects comparisons, we isolated the mean amplitudes of the electrodes comprising each significant cluster across the time range that each cluster was found. Then, using the cluster-wise average of amplitude means across the scalp, we completed a 3 X 2 X 2 ANOVA to determine differences in activation by list valence, group, and cortical region (frontal, central, and posterior). See Appendix B for full summary statistics.

Procedure.

Pretest. Participants completed a pretest to gain exposure to the format of the experimental tasks. The pretest consisted of randomized, individually presented from one of two word lists. Both lists contained ten words that were orthographically associated with neutral critical lures. During encoding, each word was presented for 2500ms, preceded by a 500ms interstimulus interval with a central fixation cross. Immediately following encoding, participants completed a recognition memory test for the previously encoded list words and the critical lures for both lists (22 total words). Test items were presented individually, and participants had to endorse each item as either old (previously encoded) or new (novel item). Participants were instructed to provide a response to the test as each item was presented. Each test trial lasted for 2500ms, after which a fixation cross was be presented for 500ms before the test program advanced to the next trial. If a response was recorded before the end of a given test trial, the trial would automatically advance.

Encoding. For each encoding block, participants indicated their level of fatigue on the KSS. Then, they encoded randomized, individually presented words from ten words lists (five associated with neutral critical lures, and five associated with negative critical lures; 80 words per block, 320 words across blocks). To ensure maintained attention to the encoding stimuli, participants were instructed to indicate whether the number of vowels in each word was even or odd as the words were presented. The duration of each encoding trial was 2500ms; the trials advanced immediately following vowel count responses, or automatically after 2500ms if no response was recorded. Following each trial, a fixation cross was shown at the center of the screen for 500ms.

Test. Immediately following the completion of each encoding block, participants completed a recognition memory test for the items presented during the most recent encoding block. The test blocks included the 80 encoded list words from the most recent encoding block, and the eight critical words for each of the respective lists. In addition, each test block began with six novel words unrelated to any other stimuli to eliminate potentially inflated response times for true test items at the beginning of the test. Each test block thus included 94 words, for a total of 376 words presented across all four test blocks. Participants were instructed to rate each word as either old or new as it was presented. Duration of each test trial was 2500ms. Each trial advanced following an old or new response from the participant, or automatically after 2500ms if no response was recorded. Following each trial, a fixation cross was shown at the center of the screen for 500ms.

Finally, participants completed the CESD and a demographic questionnaire. The CESD was completed at the end of the session to reduce the potential introduction of negative bias following the completion of a measure of depressive symptoms.

Behavioral Results.

We first assessed accuracy of the vowel count responses to each item during encoding. We found neither a significant main effect of group, $F(1,88) = .31$, $p = .58$, $\mu_p^2 = .003$, or list valence, $F(1,88) = .14$, $p = .71$, $\mu_p^2 = .002$. The interaction between the factors was not significant, $F(1,88) = .05$, $p = .83$, $\mu_p^2 < .001$.

Next, we assessed the hit rates for the list words presented during encoding (Figure 3; Table 2). We found a significant main effect of group, such that the hit rate for the DP group was significantly greater than the HC group, $F(1,88) = 10.21$, $p = .002$, $\mu_p^2 = .104$. We did not find a significant effect of list valence, $F(1,88) = .11$, $\mu_p^2 < .001$, nor a significant interaction between the factors, $F(1,88) = .07$, $\mu_p^2 < .001$. Next, we tested the false alarm rates for lure words presented only during the test blocks (Figure 3; Table 2). We found significantly fewer false alarms in the HC group than the DP group, $F(1,88) = 4.07$, $p = .04$, $\mu_p^2 = .044$. In addition, we found significantly greater false alarms for neutral lures than negative lures, $F(1,88) = 9.92$, $p = .002$, $\mu_p^2 = .20$. The interaction between the factors was not significant, $F(1,88) = .83$, $p = .36$, $\mu_p^2 = .01$.

Next, we assessed response times at test for list items (hit and miss responses). First, we found significantly quicker response times for hits than miss responses, $F(1,176) = 45.93$, $p < .001$, $\mu_p^2 = .21$ (Figure 4; Table 2). We also found quicker response times for the DP group than the HC group, $F(1,176) = 17.01$, $p < .001$, $\mu_p^2 = .09$. The main effect of list valence, however, did not have a significant effect of response times, $F(1,176) = .145$, $p = .70$, $\mu_p^2 < .001$. None of the second order interactions between the factors were significant, F's < 1.8, p's > .19. In addition, the interaction term between all three factors did not reach significance, $F(1,176) = .003$, $p = .95$, $\mu_p^2 < .001$.

We next tested response times during the test for responses to critical lures (false alarms and correct rejections). We found a main effect of group, with significantly quicker response times to critical lures by the DP group than the HC group, $F(1,174) = 14.03$, $p < .001$, $\mu_p^2 = .07$ (Figure 4; Table 2). We also found longer response times for correct rejections than false alarms, $F(1,174) = 32.68$, $p < .001$, $\mu_p^2 = .16$. The main effect of list valence did not reach statistical significance, $F(1,174) = .50$, $p = .48$, $\mu_p^2 = .002$. In addition, the second-order interaction terms between each of the main effects did not reach significance, F's < 2.1, p's > .15. Finally, the interaction term including all three factors did not reach significance, $F(1,174) = .09$, $p = .77$, $\mu_p^2 < .001$.

We conducted a 2 X 2 ANOVA to assess the scores on the KSS, with group and block order as the main effects. We also included a within-subjects error term to control for within-subject variability across blocks. We found significantly greater scores (more fatigue) in the DP group than the HC group, $F(1, 177) = 30.21$, $p < .001$, $\mu_p^2 = .15$. However, neither the main effect of block order, $F(1, 177) = .2$, $p = .66$, $\mu_p^2 = .001$, nor the interaction between the factors, $F(1, 177) = .91$, $p = .34$, $\mu_p^2 = .005$ were significant. This reported fatigue, however, did not appear to negatively impact overall accuracy between groups. Furthermore, the more fatigued DP group responded to test items more quickly than the HC group.

Table 2. Experiment 2 – Significant Behavioral Results

Exp. 2 - Significant Behavioral Results		Healthy Control Group		Depressive Group	
	Valence	Mean	SD	Mean	SD
List Word Hit Rate	Neutral	0.68	0.10	0.75	0.12
	Negative	0.68	0.12	0.76	0.11
Critical Lure FA Rate	Neutral	0.35	0.17	0.31	0.19
	Negative	0.50	0.18	0.39	0.22
Hit RT	Neutral	0.93	0.13	0.82	0.12
	Negative	0.92	0.14	0.82	0.12
Miss RT	Neutral	1.04	0.16	0.99	0.11
	Negative	1.03	0.16	0.97	0.12
False Alarm RT	Neutral	0.97	0.18	0.81	0.15
	Negative	0.94	0.17	0.84	0.17
Correct Rejection RT	Neutral	1.08	0.18	1.01	0.18
	Negative	1.04	0.15	0.99	0.10

Figure 3. Experiment 2: Mean Hit and False Alarm Rates

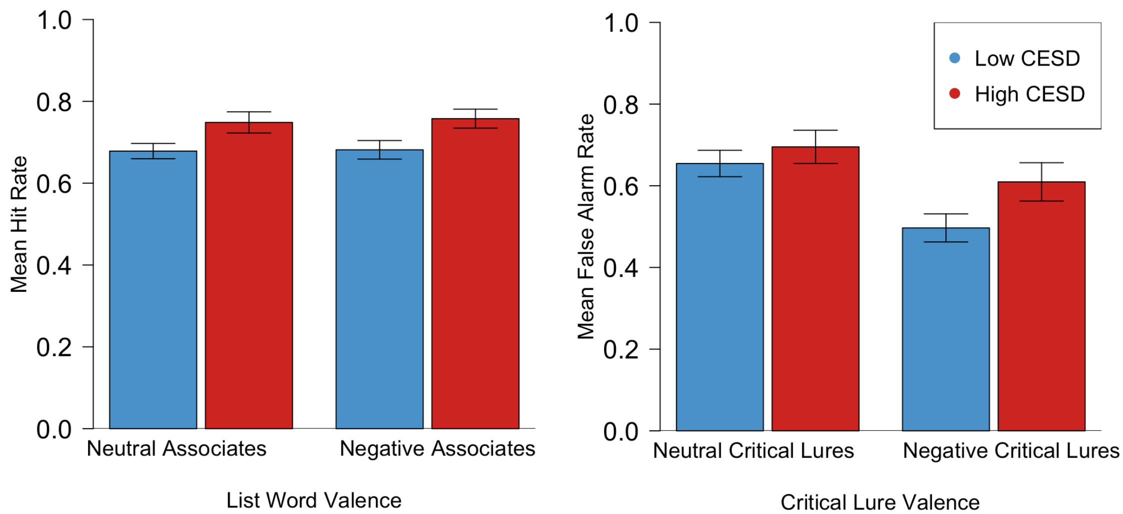

Figure 3. On the left: mean hit rates (±1SE) in Experiment 2 to list words by group and valence of the associated critical lure. On the right: mean false alarm rates (±1SE) to critical lures by group and valence.

Figure 4. Experiment 2: Mean Response Times

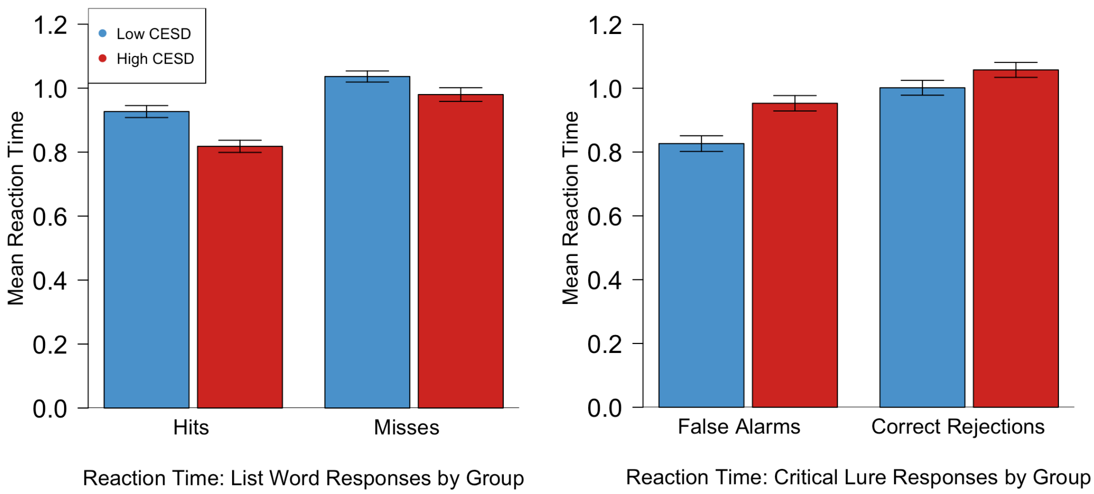

Figure 4. Mean response times in Experiment 2 for each response type by group. On the left: mean response time (±1SE) by group for hits and misses. On the right: mean response time (±1SE) by group for false alarms and correct rejections.

EEG Results.

Pointwise Non-Parametric Randomized Permutation Analysis

For the present study, we were interested in significant clusters arising in the medial frontal and left parietal cortex in accordance patterns of cortical activation for attention processing and memory retrieval. The permutation tests compared cortical activation within-subjects, analyzing the differences in reactivity for responses to list words (hits and misses) and to critical lures (false alarms and correct rejections). For both tests, all electrode locations in the 64-channel montage were included in the analysis, testing epochs from 200ms pre-stimulus onset to 1500ms post-stimulus onset. Thus, the results of the test would reflect patterns of cortical activity associated with memory retrieval, including the N400, LPC, and post-retrieval monitoring.

For the analysis of list word responses, the permutation test showed two significant clusters of activation differences (Figure 5). The first cluster was found from 316 to 676ms post-stimulus onset and distributed through all three regions of the brain over left-central electrode locations. Data points comprising the first cluster are dark red, indicating more positive amplitudes for hit responses than misses during the respective interval. The second cluster was more widespread across the duration of the tested epochs, with significant activation differences found from 716ms to the end of the epochs at 1500ms post-stimulus onset. Widespread activation differences were found across the frontal and central regions, while significant differences in the posterior region were mainly isolated to lateral electrode locations. The yellow color of the data points in the second cluster indicates more positive activation for the miss responses than hit responses.

In the analysis of critical lures responses, we found three significant clusters of activity differences, each corresponding to a critical memory retrieval component (Figure 6). The first cluster was found from 296-536ms post-stimulus onset, with widespread activity differences in the frontal region and left-lateralized differences in the central region. Our findings are consistent with the locations and timing of the N400, with the color of the clusters indicating a more positive N400 to false alarms (and more negative, i.e. stronger, N400 to correct rejections). The second cluster was found from 652-964ms post-stimulus onset, with the cluster located in left-lateralized central and posterior regions. This result is consistent with the LPC, and the color indicates more positive amplitudes for correct rejections than false alarms. The third cluster was found from 1230-1500ms post-stimulus onset and indicates more positive amplitudes for false alarms than correct rejections. Again, the locations of the significant clusters mirror prior research on post-retrieval monitoring, with bilateral activation primarily in the central and posterior regions.

Figure 5. Experiment 2: Permutation Results – False Alarms and Correct Rejections

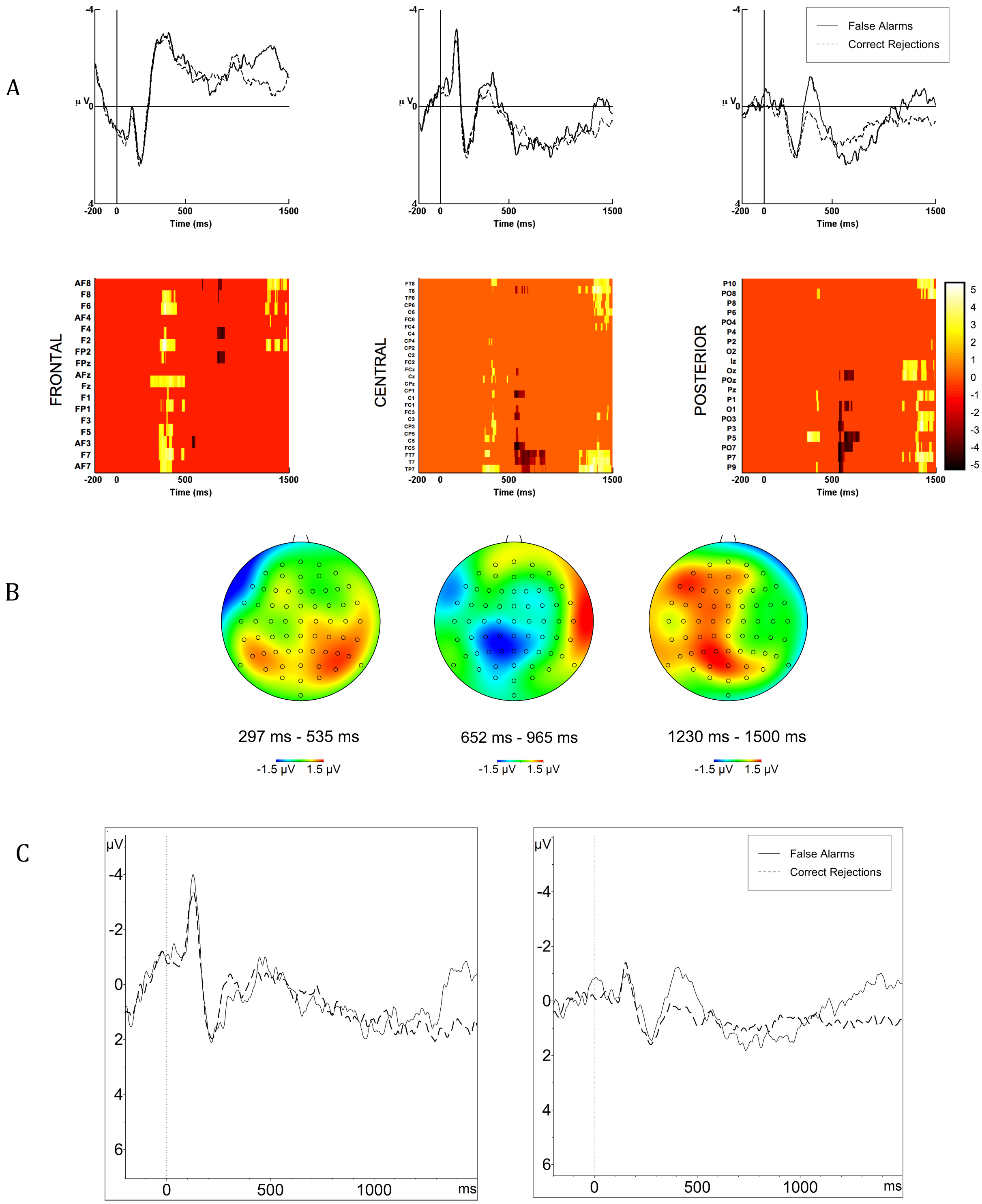

Figure 5. A) Within-subject event-related potential (ERP) differences by critical lure responses (false alarms and correct rejections). The solid lines indicate false alarms, and the dashed lines indicate correct rejections. The three ERP waveforms from left to right represent frontal, central, and posterior brain regions, respectively. The solid black vertical line at time point zero represents stimulus onset. The waveforms are derived from electrodes included in the clusters of statistical significance from the pointwise permutation testing. The color-blocked panels below the waveforms show clusters of significant activation differences between false alarms and correct rejections. Clusters are organized by brain region (frontal, central, and posterior). Electrodes within each cluster are organized from left lateralization (at the bottom of the y-axis) to right lateralization (at the top of the y-axis). The orange background indicates no significant cluster, light clusters indicate more positive activation for correct rejections, and dark clusters indicate more positive activation for false alarms. B) Topographical maps showing the locations of activation differences, false alarms minus correct rejections. Each map corresponds to the timing of each of the significant clusters from the permutation test. Red shading indicates greater activation for correct rejections, and blue shading indicates greater activation for false alarms. C) Waveforms derived from neighboring electrode locations within clusters to show the difference in response for false alarm and correct rejection responses in frontal and left parietal locations, respectively. The waveform on the left is derived from the grand average across participants at locations AF3, AF7, F3, F5, and F7. The waveform on the right is derived from the grand average across participants at locations CP5, P3, P5, P7, and TP7.

Figure 6. Experiment 2: Permutation Results – Hits and Misses

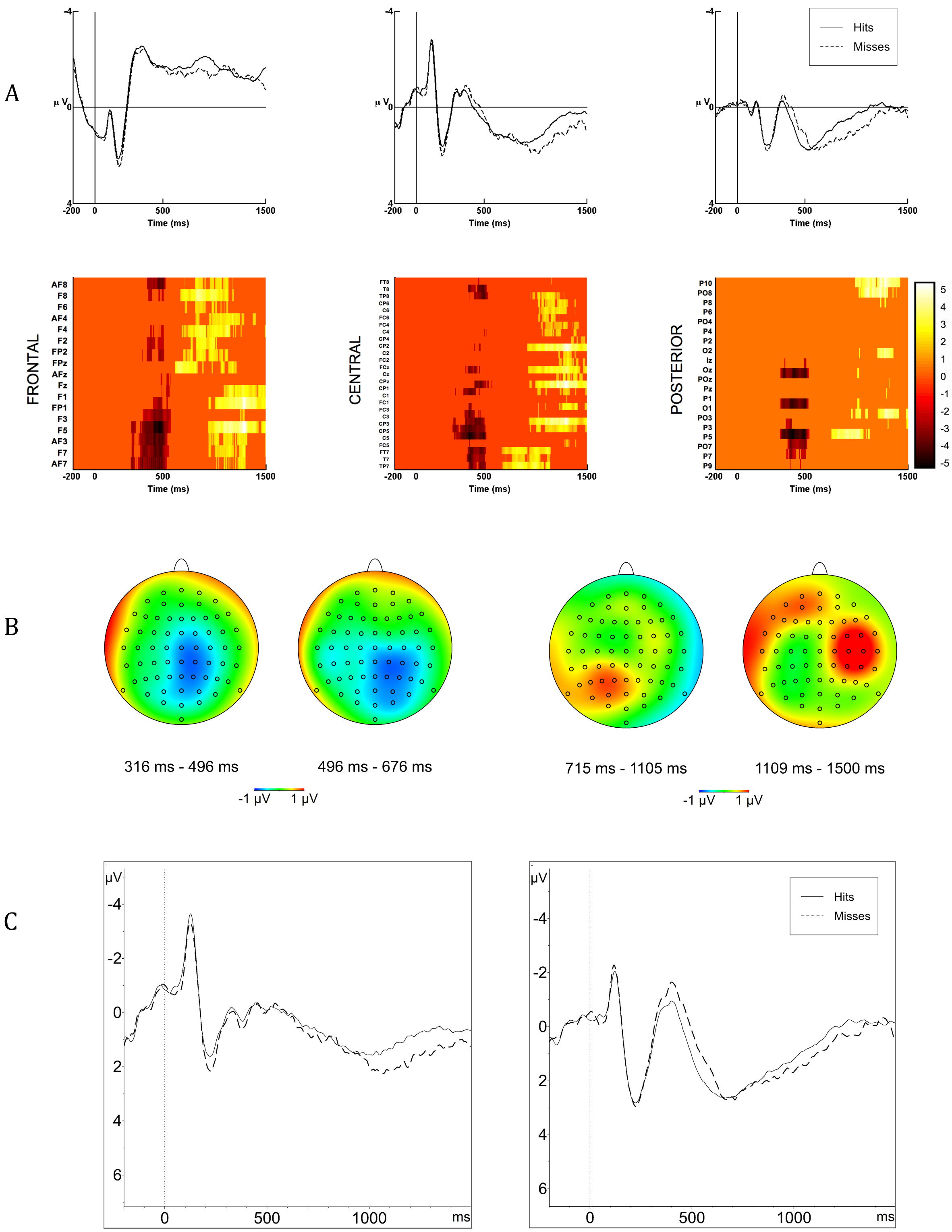

Figure 6. A) Within-subject event-related potential (ERP) differences by list word responses (hits and misses). The solid lines indicate hits, and the dashed lines indicate misses. The three ERP waveforms from left to right represent frontal, central, and posterior brain regions, respectively. The solid black vertical line at time point zero represents stimulus onset. The waveforms are derived from electrodes included in the clusters of statistical significance from the pointwise permutation testing. The color-blocked panels below the waveforms show clusters of significant activation differences between false alarms and correct rejections. Clusters are organized by brain region (frontal, central, and posterior). Electrodes within each cluster are organized from left lateralization (at the bottom of the y-axis) to right lateralization (at the top of the y-axis). The orange background indicates no significant cluster, light clusters indicate more positive activation for correct rejections, and dark clusters indicate more positive activation for false alarms. B) Topographical maps showing the locations of activation differences, miss minus hits. Each map corresponds to the timing of each of the significant clusters from the permutation test. Red shading indicates greater activation for hits, and blue shading indicates greater activation for misses. C) Waveforms derived from neighboring electrode locations within clusters to show the difference in response for hit and miss responses in frontal and central locations, respectively. The waveform on the left is derived from the grand average across participants at locations AF3, F3, F5, and F7. The waveform on the right is derived from the grand average across participants at locations C1, C2, CP1, CP2, Cz, and CPz.

ANOVA: Cluster-wise Group and List Valence Testing

Based on the results of the permutation analyses, we subsequently conducted ANOVAs to test the effects on group, valence, and brain region for each of the significant clusters. Thus, we ran five 3 X 2 X 2 ANOVA models – one for each of the two clusters representing activation differences in list word responses, and one for each of the three clusters representing activation differences in critical lure responses. All electrodes comprising each respective cluster were included to create the grand averages used in testing; for each test, the electrodes contributing the clusters in each permutation brain region were averaged together.

In the analysis of the first (earlier) cluster found in the list word permutation test, we found significantly more positive activation in the HC group than the DP group, $F(1, 540) = 52.58$, $p < .001$, $\mu_p^2 = .105$ (Table 3). In addition, we found a significant main effect of brain region, $F(2, 540) = 15.53$, $p < .001$, $\mu_p^2 = .05$, such that amplitude was most negative in the frontal region, becoming more positive in the central region, and most positive in the posterior region. We did not, however, find a significant main effect of list valence, $F(1, 540) = .04$, $p = .85$, $\mu_p^2 < .001$. The interaction of group and brain region narrowly missed significance, $F(2, 540) = 2.85$, $p = .06$, $\mu_p^2 = .01$. In this interaction, both groups showed increasingly positive amplitude from the frontal to the posterior region. However, the range of mean amplitude values was greater in the DP group than in the HC group. The remaining interaction terms for valence and group, $F(1, 540) = .75$, $p = .39$, $\mu_p^2 = .001$, valence and brain region, $F(2, 540) = .22$, $p = .80$, $\mu_p^2 < .001$, and all three main factors, $F(2, 540) = .58$, $p = .56$, $\mu_p^2 < .001$, did not reach significance.

In the analysis of the second list words cluster, we found a significant main effect of brain region, $F(2, 540) = 29.19$, $p < .001$, $\mu_p^2 = .10$ (Table 3). We found the most positive amplitude in the frontal region, becoming more negative in the central region, and the most negative in the posterior region. We did not find a main effect of list valence, $F(1, 540) = .08$, $p = .78$, $\mu_p^2 < .001$, or group, $F(1, 540) = 1.06$, $p = .30$, $\mu_p^2 = .002$. Further, none of the interaction terms were statistically significant; list valence and group: $F(1, 540) = 1.15$, $p = .28$, $\mu_p^2 = .002$; list valence and brain region: $F(2, 540) = .06$, $p = .94$, $\mu_p^2 < .001$; group and brain region: $F(2, 540) = .63$, $p = .53$, $\mu_p^2 = .002$; list valence, group, and brain region: $F(2, 540) = .13$, $p = .88$, $\mu_p^2 < .001$.

Next, we analyzed the first cluster from the critical lure permutation test. We found significantly more negative mean amplitudes in the DP group than the HC group, $F(1, 525) = 16.80$, $p < .001$, $\mu_p^2 = .03$ (Figure 7; Table 4). We also found a significant main effect of brain region, $F(2, 525) = 4.16$, $p = .02$, $\mu_p^2 = .02$, with the most negative mean amplitudes in the frontal region, and increasingly positive amplitudes in the central and posterior regions, respectively. The main effect of list valence was not significant, $F(1, 525) = 1.35$, $p = .25$, $\mu_p^2 = .003$. Interestingly, the interaction between group and brain region did not reach significance, $F(2, 525) = 1.69$, $p = .19$, $\mu_p^2 = .006$; both groups showed similar distributions of activation amplitude from the frontal to posterior regions, with more negative amplitude overall in the DP group than the HC group. None of the remaining interaction terms were significant; list valence and group: $F(1, 525) = .05$, $p = .83$, $\mu_p^2 < .001$; list valence and brain region: $F(2, 525) = 1.11$, $p = .33$, $\mu_p^2 = .004$; list valence, group, and brain region: $F(2, 525) = .25$, $p = .78$, $\mu_p^2 < .001$.

For the analysis of the second critical lure cluster, we found that the main effect of group narrowly missed significance, $F(1, 525) = 3.71$, $p = .055$, $\mu_p^2 = .007$, with more positive mean activation amplitude in the HC group than the DP group (Figure 8; Table 4). Neither the main effect of list valence, $F(1,525) = .17$, $p = .68$, $\mu_p^2 < .001$, nor brain region, $F(2, 525) = 1.12$, $p = .33$, $\mu_p^2 = .004$, reached significance. We also found that the interaction between list valence and group narrowly missed significance, $F(1, 525) = 3.56$, $p = .06$, $\mu_p^2 = .007$. Here, we saw that both groups had nearly identical mean amplitudes for neutral critical lures. However, the mean amplitude for the HC group was greater than the DP group in response to negative critical lures. None of the remaining interaction terms reached statistical significance; list valence and brain region: $F(2, 525) = .59$, $p = .55$, $\mu_p^2 = .002$; group and brain region: $F(2, 525) = .24$, $p = .80$, $\mu_p^2 < .001$; list valence, group, and brain region: $F(2, 525) = 2.08$, $p = .13$, $\mu_p^2 = .008$.

Finally, we analyzed the third cluster from the critical lure permutation tests. We found a significant main effect of brain region, $F(2, 525) = 7.28$, $p < .001$, $\mu_p^2 = .03$, with mean activation amplitudes most positive in the frontal region, sequentially decreasing in the central and posterior regions, respectively. The main effects of group, $F(1, 525) = 1.10$, $p = .30$, $\mu_p^2 = .002$, and list valence, $F(1, 525) = .74$, $p = .39$, $\mu_p^2 = .001$, were not significant. In addition, none of the interaction terms reached statistical significance; list valence and group: $F(1, 525) = .81$, $p = .37$, $\mu_p^2 = .002$; list valence and brain region: $F(2, 525) = .46$, $p = .63$, $\mu_p^2 = .002$; group and brain region: $F(2, 525) = .94$, $p = .39$, $\mu_p^2 = .004$; list valence, group, and brain region, $F(2, 525) = .50$, $p = .61$, $\mu_p^2 = .002$.

Table 3. Experiment 2 – EEG Mean Amplitudes – Hit/Miss Clusters

Exp. 2 EEG Mean Amplitudes (μV)	Brain Region	Group	Valence	Mean	Std. Dev.
Hit/Miss Cluster 1 (316-676ms)	Frontal	HC	Neutral	0.72	2.19
			Negative	0.23	1.95
		DP	Neutral	-1.01	2.10
			Negative	-0.84	1.85
	Central	HC	Neutral	0.98	1.53
			Negative	0.90	1.31
		DP	Neutral	-0.29	1.37
			Negative	-0.18	1.52
	Posterior	HC	Neutral	1.00	1.73
			Negative	1.11	1.69
		DP	Neutral	0.47	1.55
			Negative	0.48	1.37
Hit/Miss Cluster 2 (716-1500ms)	Frontal	HC	Neutral	1.21	1.50
			Negative	1.08	1.39
		DP	Neutral	1.31	1.41
			Negative	1.48	1.61
	Central	HC	Neutral	1.03	1.20
			Negative	0.84	1.20
		DP	Neutral	1.00	1.23
			Negative	1.12	1.43
	Posterior	HC	Neutral	0.37	0.79
			Negative	0.26	1.21
		DP	Neutral	0.29	0.85
			Negative	0.26	0.96

Table 4. Experiment 2 – EEG Mean Amplitudes – False Alarm/Correct Rejection Clusters

Exp. 2 EEG Mean Amplitudes (μV)	Brain Region	Group	Valence	Mean	Std. Dev.
FA/CR Cluster 1 (296 – 536ms)	Frontal	HC	Neutral	-0.06	3.45
			Negative	1.18	3.45
		DP	Neutral	-1.43	4.57
			Negative	-0.88	3.63
	Central	HC	Neutral	0.53	3.47
			Negative	0.80	2.82
		DP	Neutral	-0.88	3.33
			Negative	-0.60	2.97
	Posterior	HC	Neutral	1.07	3.29
			Negative	0.77	3.25
		DP	Neutral	0.47	3.36
			Negative	0.46	2.11
FA/CR Cluster 2 (652-964ms)	Frontal	HC	Neutral	0.74	4.56
			Negative	1.83	4.30
		DP	Neutral	1.68	3.65
			Negative	0.27	2.29
	Central	HC	Neutral	1.53	4.20
			Negative	1.96	3.03
		DP	Neutral	1.22	3.22
			Negative	0.71	2.10
	Posterior	HC	Neutral	0.91	3.49
			Negative	1.31	2.12
		DP	Neutral	0.24	2.40
			Negative	0.92	2.54

Figure 7. Experiment 2: FA/CR Cluster 1 – Activation by Group and Valence

A

B

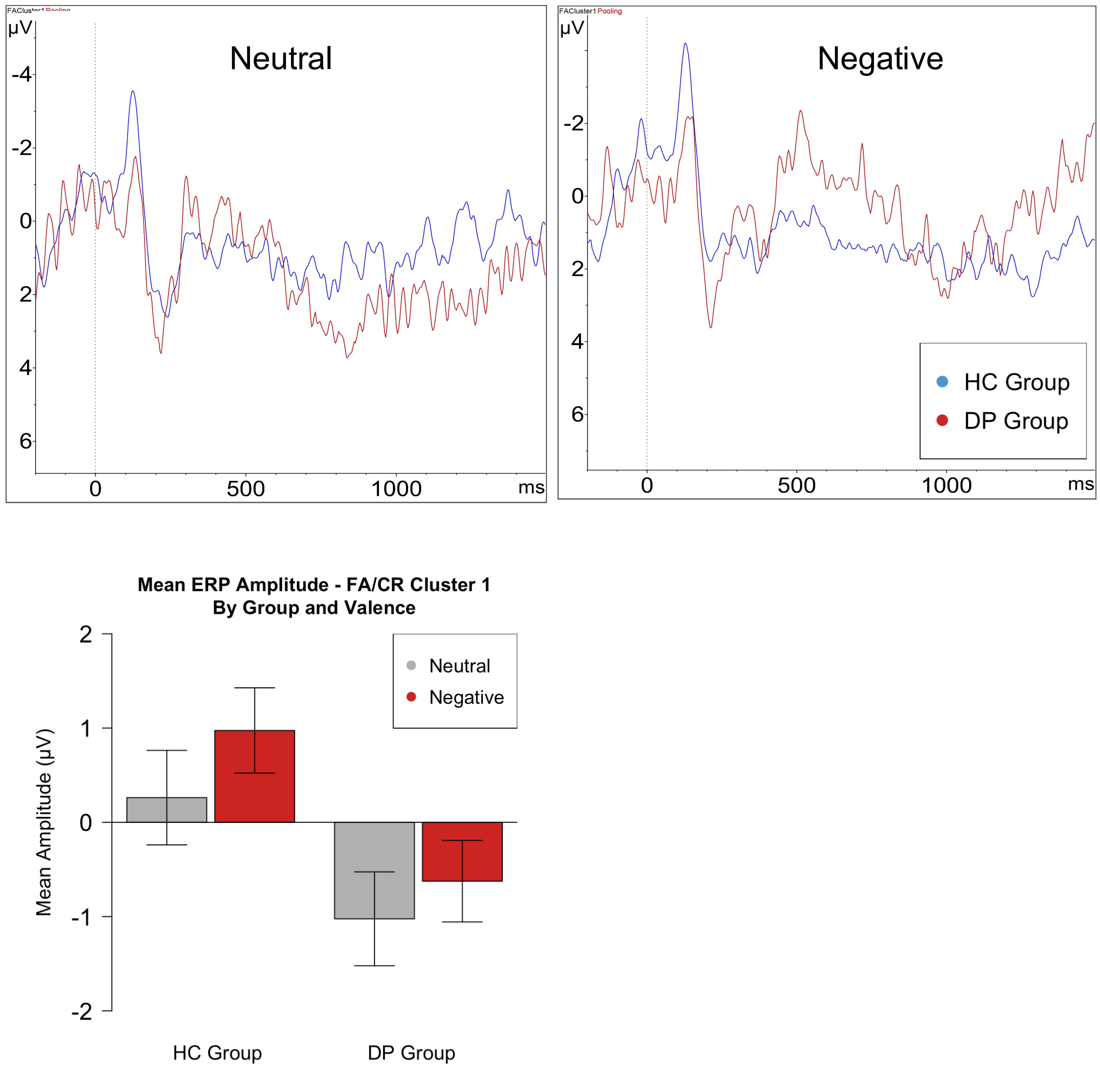

Figure 7. A) Waveforms derived from neighboring electrode locations within the first cluster of significant activation from the permutation test of activation for false alarms and correct rejections. The waveform on the left is derived from the grand average across participants at locations AF3, F3, F5, and F7 to show differences in early frontal responses, which reached statistical significance in the permutation testing. The waveform on the left shows differences in responses to neutral critical lures by group, with the healthy control group in blue and the depressive symptom group in red. The waveforms on the right show difference in response to negative critical lures by group, with the healthy control group in blue and the depressive symptom group in red. B) Mean ERP amplitude (±1SE) averaged across all electrodes that reached significance in the first cluster of the false alarm and correct rejection permutation test. Means are presented by group and valence (neutral and negative).

Figure 8. Experiment 2: FA/CR Cluster 2 – Activation by Group and Valence

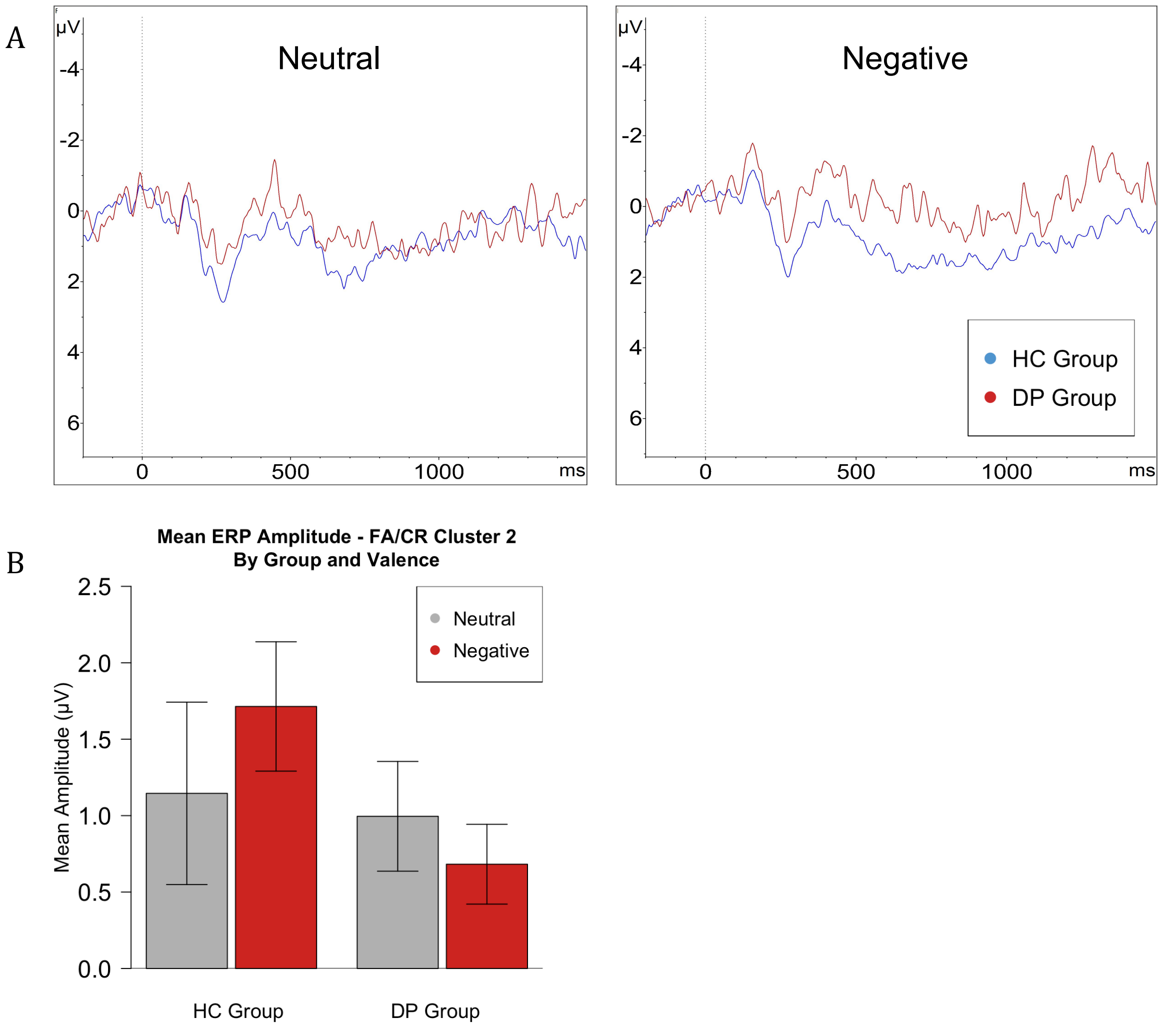

Figure 8. A) Waveforms derived from neighboring electrode locations within the first cluster of significant activation from the permutation test of activation for false alarms and correct rejections. The waveform on the left is derived from the grand average across participants at locations C1, C2, CP1, CP2, Cz, and CPz to show differences in late central parietal responses, which reached statistical significance in the permutation testing. The waveform on the left shows differences in responses to neutral critical lures by group, with the healthy control group in blue and the depressive symptom group in red. The waveforms on the right show difference in response to negative critical lures by group, with the healthy control group in blue and the depressive symptom group in red. B) Mean ERP amplitude (±1SE) averaged across all electrodes that reached significance in the first cluster of the false alarm and correct rejection permutation test. Means are presented by group and valence (neutral and negative).

Discussion

In Experiment 2, we collected EEG data during an orthographic recognition memory test to determine differences in retrieval-related activation patterns between individuals with depressive symptoms and healthy controls. Our behavioral findings showed that the depressive symptom group had greater overall hit rates and false alarm rates than the control group. Thus, the depressive symptom group had a significantly more liberal criterion for endorsing test items as old; although false alarm rates were higher, the greater hit rate meant that overall accuracy between groups was similar. Furthermore, the depressive symptom group exhibited quicker response times for all test items than the control group while not negatively affecting their overall accuracy.

Our permutation testing showed activation patterns consistent with findings from the memory retrieval literature. In the comparison between hits and misses, the greater positive amplitudes for hits in the first cluster falls in a similar time range as the late positive component, which often shows greater amplitude for correct recognition of previously learned items. The second cluster showed widespread late positive activation in the frontal and central brain regions, with greater activation amplitude for misses than hit responses. This activity supports the occurrence of post-retrieval monitoring, which commonly occurs in frontal cortical areas after making an error in memory retrieval. Similarly, the significant clusters found in the comparison of correct rejections and false alarms to critical lures follow the components of the dual process theory of memory retrieval. The first and third clusters show greater amplitudes for false alarms than correct rejections, with the former occurring in the same time and cortical regions as the N400, and the latter occurring in broader areas to reflect post-retrieval monitoring. Finally, the middle cluster showed greater amplitude for correct rejections than false alarms, focused in left central and left posterior regions in accordance with the late positive component.

The statistical tests for group and valence at the locations and time ranges indicated by the permutation tests more strongly illuminated the differences between the depressive and control groups. First, we found that the control group had more positive amplitudes in the first cluster region for hit responses than the depressive symptom group. This finding supports the idea of confidence correlating with EEG amplitude, as found in the behavioral results of both experiments. Confidence ratings from Experiment 1 were reduced in the depressive group relative to controls for hit responses, which would suggest that sample might have also shown reduced cortical activation for hits. Further, Experiment 2 showed that the control group was more conservative in endorsing items as hits than the depressive group. This heightened scrutiny of hit endorsements may suggest that the control group was more confident in their hit endorsements than the more liberal endorsements from the depressive symptom group.

More critically, we found evidence that the control group and the depressive group may respond differently to false test items. First, the significantly greater activation by the depressive symptom group than the controls in the first critical lure cluster suggests different familiarity-related responses. The time and location of the first cluster mirror the N400 as found in non-permutation analyses, which in memory retrieval paradigms is associated with familiarity-based retrieval. In contrast, the control group showed greater activation than the depressive group in the second cluster, which mirrors the recollection-related LPC. Although the second finding narrowly missed significance, the potential differences in retrieval strategy are supported by behavioral data. The depressive group may utilize a more familiarity-dependent retrieval strategy; retrieving items on a generic level would show greater early frontal activation and allow for quicker responses to each item at the expense of greater false alarm rates. In contract, the control group may utilize a more recollection-dependent strategy. Endorsing items as old given the presence of a stronger, more detailed memory for each item would result in greater late positive retrieval activation and longer response times during each test trial.

The present study has two major limitations. First, participants were not presented with entirely novel lists during test; the only novel information they viewed were the critical lures associated with the list words they already encoded. To form a greater understanding of the difference in response to novel non-target stimuli, additional research is needed which includes novel list words and critical lures independent of anything shown during encoding. In addition, our study assumed the presence of negative attention bias in the individuals with depressive symptoms. Although negative attention bias is very often comorbid with depressive symptoms, its presence is not guaranteed, and the prevalence of the bias may vary. Therefore, a direct measure of negative attention bias would allow for stronger conclusions to be drawn about the effect of bias on memory distortion.

Together, the two experiments in this study provide evidence that individuals with depressive symptoms and healthy individuals may have slightly different memory retrieval strategies for emotional false information. However, additional research is necessary to create a greater understand of the potential differences in retrieval. In addition to directly measuring negative attention bias, more information is needed to determine the mechanism of the differences observed in the present study. For example, differing retrieval strategies may not be intrinsic to depressive symptoms and may also be elicited in control participants. Addressing further questions about the relationship between negative attention bias and depressive symptoms will be critical for continuing to elucidate the nature of retrieval strategies in a memory distortion framework.

Appendix A. Word lists used in Experiment 1. Critical lures are entered at the beginning of each list in bold font, and list words follow in standard font.

Pretest Items:

Sail	**Push**
Mail	Posh
Rail	Pus
Hail	Rush
Gail	Mush
Nail	Lush
Quail	Brush
Stale	Plush
Scale	Bush
Whale	Hush
Veil	Such

Encoding/Test Items:

Lists with Negative Critical Lures

Rape	**Bitch**	**Hell**	**Trash**	**Whore**	**Pain**	**Slut**	**Fear**
Cape	Ditch	Bell	Brash	Chore	Pail	Slug	Feat
Nape	Hitch	Dell	Cash	Bore	Paid	Slum	Feed
Tape	Batch	Fell	Clash	Wore	Pair	Slur	Tear
Ripe	Pitch	Jell	Flash	More	Main	Slot	Wear
Rope	Itch	Sell	Slash	Tore	Rain	Slue	Pear
Race	Botch	Tell	Smash	Pore	Chain	Shut	Dear
Rapt	Mitch	Hall	Splash	Sore	Stain	Slit	Bear
Rake	Butch	Hill	Stash	Horn	Train	Smut	Year
Rare	Birch	Hull	Dash	Shore	Vein	Glut	Shear
Raze	Witch	Shell	Bash	Lore	Pan	Scut	Flare

Lists with Neutral Critical Lures

Hook	**Peach**	**Rink**	**Digit**	**Shave**	**Vest**	**Park**	**Cord**
Book	Beach	Link	Widget	Slave	Vast	Bark	Card
Look	Leach	Mink	Midget	Stave	West	Dark	Curd
Cook	Teach	Sink	Bridget	Shove	Rest	Hark	Core
Nook	Reach	Wink	Fidget	Share	Test	Lark	Cold
Rook	Poach	Pink	Divot	Have	Zest	Mark	Lord
Took	Peak	Rank	Divvy	Shade	Nest	Nark	Ford
Hick	Perch	Risk	Dimwit	Sake	Lest	Pack	Word
Honk	Peace	Blink	Digest	Shale	Crest	Perk	Gourd
Hood	Preach	Rick	Gidget	Shame	Chest	Pork	Court
Hoof	Peal	Fink	Dig	Shape	Jest	Spark	Sword

Lists with Positive Critical Lures

Merry	**King**	**Cheer**	**Gold**	**Thrill**	**Puppy**	**Treat**	**Cute**
Sherry	Bing	Hear	Hold	Frill	Peppy	Trees	Cite
Berry	Ding	Beer	Good	Drill	Yuppie	Tweet	Cube
Cherry	Ping	Near	Goal	Grill	Poppy	Tread	Chute
Ferry	Ring	Spear	Told	Trill	Pappy	Head	Flute
Dairy	Sing	Clear	Mold	April	Preppy	Peat	Mute
Airy	Ting	Veer	Bold	Shrill	Putty	Seat	Newt
Bury	Wing	Deer	Fold	Still	Hubby	Meat	Hoot
Hairy	Zing	Gear	Guild	Will	Puffy	Meet	Shoot
Very	Cling	Jeer	Sold	Mill	Pupil	Pleat	Loot
Wary	Bring	Steer	Cold	Trail	Poopy	Beet	Root

Appendix B. Summary Statistics

Table 1.

Exp. 1 Demographics	Healthy Control Group	Depressive Group
N	32	32
Age: Mean (SD)	19.4 (1.24)	20.3 (2.32)
Female	24	23
White	14	14
Hispanic	13	8
CESD Score: Mean (SD)	9.37 (3.24)	28.75 (6.98)

Table 2.

Exp. 1 Behavioral Results

	Valence	Healthy Control Group Mean	SD	Depressive Group Mean	SD
List Word Hit Rate	Positive	0.74	0.12	0.69	0.11
	Neutral	0.71	0.12	0.74	0.12
	Negative	0.72	0.10	0.68	0.11
Critical Lure False Alarm Rate	Positive	0.25	0.20	0.34	0.23
	Neutral	0.54	0.25	0.60	0.24
	Negative	0.26	0.21	0.27	0.21
Hit Response Time (RT)	Positive	1.47	0.24	1.49	0.36
	Neutral	1.45	0.21	1.50	0.39
	Negative	1.52	0.31	1.52	0.44
Miss RT	Positive	1.62	0.32	1.70	0.59
	Neutral	1.61	0.30	1.76	0.49
	Negative	1.63	0.38	1.76	0.67
False Alarm (FA) RT	Positive	1.49	0.59	1.80	0.86
	Neutral	1.55	0.46	1.45	0.52
	Negative	1.53	0.55	1.68	1.05
Correct Rejection (CR) RT	Positive	1.49	0.29	1.51	0.35
	Neutral	1.66	0.42	1.89	0.69
	Negative	1.62	0.39	1.64	0.34
Hit Confidence Rating	Positive	3.84	0.66	3.77	0.46
	Neutral	4.20	0.64	4.18	0.45
	Negative	3.80	0.64	3.71	0.55
Miss Confidence Rating	Positive	3.30	0.74	3.04	0.70
	Neutral	3.28	0.68	3.04	0.75
	Negative	3.35	0.59	3.14	0.79
FA Confidence Rating	Positive	3.32	1.09	3.48	0.96
	Neutral	3.57	0.87	3.81	0.90
	Negative	3.25	1.09	3.55	0.96
CR Confidence Rating	Positive	3.61	0.84	3.32	0.80
	Neutral	3.54	0.78	2.94	1.02
	Negative	4.04	0.79	4.01	0.70

Table 3.

Exp. 2 Demographics	Healthy Control Group	Depressive Group
N	21	25
Age: Mean (SD)	19.9 (1.84)	19.5 (1.07)
Female	12	19
White	10	15
Hispanic	9	5
CESD Score: Mean (SD)	8.06 (3.76)	22.15 (6.61)

Table 4.

Exp. 2 Behavioral Results		Healthy Control Group		Depressive Group	
	Valence	Mean	SD	Mean	SD
List Word Hit Rate	Neutral	0.68	0.10	0.75	0.12
	Negative	0.68	0.12	0.76	0.11
Critical Lure FA Rate	Neutral	0.35	0.17	0.31	0.19
	Negative	0.50	0.18	0.39	0.22
Hit RT	Neutral	0.93	0.13	0.82	0.12
	Negative	0.92	0.14	0.82	0.12
Miss RT	Neutral	1.04	0.16	0.99	0.11
	Negative	1.03	0.16	0.97	0.12
False Alarm RT	Neutral	0.97	0.18	0.81	0.15
	Negative	0.94	0.17	0.84	0.17
Correct Rejection RT	Neutral	1.08	0.18	1.01	0.18
	Negative	1.04	0.15	0.99	0.10
KSS Score, First Block		4.38	1.50	6.04	1.57
KSS Score, Second Block		4.52	1.69	6.20	1.44
KSS Score, Third Block		4.86	1.68	5.96	1.71
KSS Score, Fourth Block		4.60	1.88	5.69	1.69

Table 5.

Exp. 2 EEG Mean Amplitudes (mV)	Brain Region	Group	Valence	Mean	Std. Dev.
FA/CR Cluster 1 (296 – 536ms)	Frontal	HC	Neutral	-0.06	3.45
			Negative	1.18	3.45
		DP	Neutral	-1.43	4.57
			Negative	-0.88	3.63
	Central	HC	Neutral	0.53	3.47
			Negative	0.80	2.82
		DP	Neutral	-0.88	3.33
			Negative	-0.60	2.97
	Posterior	HC	Neutral	1.07	3.29
			Negative	0.77	3.25
		DP	Neutral	0.47	3.36
			Negative	0.46	2.11
FA/CR Cluster 2 (652-964ms)	Frontal	HC	Neutral	0.74	4.56
			Negative	1.83	4.30
		DP	Neutral	1.68	3.65
			Negative	0.27	2.29
	Central	HC	Neutral	1.53	4.20
			Negative	1.96	3.03
		DP	Neutral	1.22	3.22
			Negative	0.71	2.10
	Posterior	HC	Neutral	0.91	3.49
			Negative	1.31	2.12
		DP	Neutral	0.24	2.40
			Negative	0.92	2.54
FA/CR Cluster 3 (12320 – 1500ms)	Frontal	HC	Neutral	1.10	4.97
			Negative	1.14	3.52
		DP	Neutral	1.99	4.48
			Negative	0.69	2.93
	Central	HC	Neutral	0.65	4.11
			Negative	0.57	2.26
		DP	Neutral	0.36	3.50
			Negative	0.09	2.14
	Posterior	HC	Neutral	0.22	3.75
			Negative	0.29	2.55
		DP	Neutral	-0.51	2.62
			Negative	-0.48	2.37

Table 6.

Exp. 2 EEG Mean Amplitudes	Brain Region	Group	Valence	Mean	Std. Dev.
Hit/Miss Cluster 1 (316-676ms)	Frontal	HC	Neutral	0.72	2.19
			Negative	0.23	1.95
		DP	Neutral	-1.01	2.10
			Negative	-0.84	1.85
	Central	HC	Neutral	0.98	1.53
			Negative	0.90	1.31
		DP	Neutral	-0.29	1.37
			Negative	-0.18	1.52
	Posterior	HC	Neutral	1.00	1.73
			Negative	1.11	1.69
		DP	Neutral	0.47	1.55
			Negative	0.48	1.37
Hit/Miss Cluster 2 (716-1500ms)	Frontal	HC	Neutral	1.21	1.50
			Negative	1.08	1.39
		DP	Neutral	1.31	1.41
			Negative	1.48	1.61
	Central	HC	Neutral	1.03	1.20
			Negative	0.84	1.20
		DP	Neutral	1.00	1.23
			Negative	1.12	1.43
	Posterior	HC	Neutral	0.37	0.79
			Negative	0.26	1.21
		DP	Neutral	0.29	0.85
			Negative	0.26	0.96

Appendix C. Word lists used in Experiment 2. Critical lures are entered at the beginning of each list in bold font, and list words follow in standard font.

Pretest Items:

Plow	**Push**
Plot	Posh
Pro	Pus
Plume	Rush
Prowl	Mush
Ply	Lush
Vow	Brush
Low	Plush
Cow	Bush
Row	Hush
Tow	Such

Encoding/Test Items:

Recency Items (at the beginning of test blocks):

Block 1	Block 2	Block 3	Block 4
Burger	Face	Sand	Veal
Whale	Quiet	Zebra	Goal
Bulb	Chili	Carbon	Pile
Locket	Number	Paper	Hind
Gel	Plank	Type	Kitten
Swing	Land	Fur	Windy

Lists with Negative Critical Lures

Fear	**Germ**	**Grief**	**Pain**	**Stress**	**Bitch**	**Hurt**	**Drown**	**Rape**	**Death**
Bear	Gem	Beef	Pan	Bless	Batch	Blurt	Crown	Rapt	Doth
Dear	Term	Brief	Chain	Chess	Birch	Dirt	Down	Race	Breath
Feat	Gram	Chief	Main	Cress	Butch	Flirt	Town	Rake	Dealt
Feed	Perm	Green	Paid	Dress	Ditch	Halt	Thrown	Rare	Dearth
Pear	Berm	Leaf	Pail	Press	Hitch	Shirt	Drawn	Raze	Ether
Tear	Glam	Strive	Pair	Strewn	Itch	Short	Drawl	Ripe	Heath
Wear	Fern	Thief	Rain	Truss	Mitch	Heart	Dawn	Rope	Sheath
Year	Stern	Relief	Stain	Trees	Pitch	Inert	Drone	Tape	Stealth

Scum	**Trash**	**Jail**	**Whore**	**Kill**	**Hell**	**Lice**	**Ache**	**Slut**	**Hate**
Chum	Bash	Hail	Horn	Bill	Fell	Lick	Act	Scut	Hats
Clump	Brash	Jade	Chore	Pill	Hall	Life	Mach	Shut	Fate
Crumb	Cash	Mail	Lore	Krill	Hill	Lime	Ace	Slat	Hare
Drum	Clash	Nail	More	Sill	Hull	Line	Each	Slit	Haste
Lump	Dash	Pail	Pore	Dill	Jell	Lite	Bake	Slot	Haze
Plum	Flash	Rail	Shore	Kilt	Sell	Mice	Lake	Slug	Late
Shim	Slash	Sail	Sore	Kiln	Shell	Rice	Cache	Slum	Mate
Skim	Stash	Wail	Wore	Gill	Tell	Splice	Chai	Slur	Plate

Lists with Neutral Critical Lures

Bowl	**Obey**	**Park**	**Shave**	**Peach**	**Time**	**Door**	**Cat**	**Vest**	**Black**
Bows	Stay	Hark	Have	Perch	Chime	Doer	Car	Jest	Back
Cowl	They	Lark	Shade	Peak	Dime	Dour	Chat	Lest	Blank
Fowl	Bane	Mark	Shale	Peal	Mime	Drawer	Caw	Nest	Block
Howl	Ode	Nark	Shame	Poach	Rhyme	Floor	Cap	Rest	Rack
Owl	Ole	Pack	Shape	Preach	Tame	Four	Can	Test	Shack
Bail	Oboe	Perk	Share	Reach	Theme	Moor	Cast	Vast	Slack
Boil	Okay	Pork	Shove	Teach	Tine	More	Cart	West	Bloke
Towel	Ordain	Spark	Stave	Peace	Tome	Roar	Cay	Zest	Plaque

Digit	**Rink**	**Arm**	**Chin**	**Cord**	**Hook**	**Hard**	**Tile**	**Knot**	**Phase**
Dig	Link	Ark	Chick	Card	Honk	Chard	Tilt	Clot	Phrase
Bridget	Mink	Art	Chill	Cold	Book	Hand	Stile	Cot	Place
Digest	Pink	Farm	Chip	Curd	Cook	Harp	While	Knit	Chase
Divot	Rank	Warm	Inch	Ford	Hood	Held	Mile	Skit	Vase
Divvy	Rick	Barn	Shin	Lord	Hoof	Herd	Tale	Snot	Pace
Fidget	Risk	Alarm	Thin	Word	Look	Shard	Time	Trot	Trace
Gidget	Sink	Armor	Chino	Core	Nook	Hired	Tire	Knife	Haze
Widget	Wink	Army	Niche	Court	Rook	Hoard	Toile	Note	Race

References

1. Wickelgren, W. A. Trace resistance and the decay of long-term memory. *J. Math. Psychol.* **9,** 418–455 (1972).
2. Murphy, G. L. & Shapiro, A. M. Forgetting of Verbatim Information in Discourse. *Mem. Cognit.* **22,** 85–94 (1994).
3. Loftus, E. F. Planting misinformation in the human mind: a 30-year investigation of the malleability of memory. *Learn. Mem.* **12,** 361–366 (2005).
4. Loftus, E. Legal and Practical Implications. *Neuroscience* **4,** 2–5 (2003).
5. Reyna, V. F., Intons-Peterson, M. & Best, D. in *Current Directions in Psychological Science* **11,** 164–168 (2002).
6. Loftus, E. F. Memory distortion and false memory creation. *Bull. Am. Acad. Psychiatry Law* **24,** 281–295 (1996).
7. Loftus, E. F. & Pickrell, J. E. The formation of false memories. *Psychiatric Annals* **25,** 720–725 (1995).
8. Loftus, E. F. & Hoffman, H. G. Misinformation and memory: the creation of new memories. *J. Exp. Psychol. Gen.* **118,** 100–104 (1989).
9. Roediger III, H. L. *et al.* Creating False Memories: Remembering Words Not Presented in Lists. *J. Exp. Psychol. Learn. Mem. Cogn.* **21,** 803–814 (1995).
10. Gallo, D. False memories and fantastic beliefs: 15 years of the DRM illusion. *Mem. Cognit.* **38,** 833–848 (2010).
11. Gotlib, I. H., Krasnoperova, E., Neubauer Yue, D. & Joormann, J. Attentional Biases for Negative Interpersonal Stimuli in Clinical Depression. *J. Abnorm. Psychol.* **113,** 127–135 (2004).

12. Clasen, P. C., Wells, T. T., Ellis, A. J. & Beevers, C. G. Attentional biases and the persistence of sad mood in major depressive disorder. *J. Abnorm. Psychol.* **122,** 74–85 (2013).

13. Schneider, F., Gur, R. C., Gur, R. E. & Muenz, L. R. Standardized mood induction with happy and sad facial expressions. *Psychiatry Res.* **51,** 19–31 (1994).

14. Knight, B. G., Maines, M. L. & Robinson, G. S. The effects of sad mood on memory in older adults: a test of the mood congruence effect. *Psychol. Aging* **17,** 653–661 (2002).

15. Rodríguez, A., Rey, B., Clemente, M., Wrzesien, M. & Alcañiz, M. Assessing brain activations associated with emotional regulation during virtual reality mood induction procedures. *Expert Syst. Appl.* **42,** 1699–1709 (2015).

16. Ellis, H. C., Thomas, R. L. & Rodriguez, I. a. Emotional mood states and memory: elaborative encoding, semantic processing, and cognitive effort. *J. Exp. Psychol. Learn. Mem. Cogn.* **10,** 470–482 (1984).

17. Paz-Alonso, P. M. & Goodman, G. S. Trauma and memory: effects of post-event misinformation, retrieval order, and retention interval. *Memory* **16,** 58–75 (2008).

18. Center for Behavioral Health Statistics and Quality. *Key substance use and mental health indicators in the United States: Results from the 2015 national survey on drug use and health*. *Key Substance Use and Mental Health Indicators in the United States: Results from the 2015 National Survey on Drug Use and Health* (2016).

19. Kensinger, E. a & Corkin, S. Effect of negative emotional content on working

memory and long-term memory. *Emotion* **3,** 378–93 (2003).

20. Kensinger, E. A. Remembering the Details: Effects of Emotion. *Emot. Rev.* **1,** 99–113 (2009).

21. Joormann, J., Teachman, B. A. & Gotlib, I. H. Sadder and Less Accurate? False Memory for Negative Material in Depression. *J. Abnorm. Psychol.* **118,** 412–417 (2009).

22. Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G. & Reyna, V. F. How Does Negative Emotion Cause False Memories? *Psychol. Sci.* **19,** 919–925 (2008).

23. Pesta, B. J., Murphy, M. D. & Sanders, R. E. Are Emotionally Charged Lures Immune to False Memory? *J. Exp. Psychol. Learn. Mem. Cogn.* **27,** 328–338 (2001).

24. Piguet, O., Connally, E., Krendl, A. C., Huot, J. R. & Corkin, S. False Memory in Aging: Effects of Emotional Valence on Word Recognition Accuracy. *Psychol. Aging* **23,** 307–314 (2008).

25. Bradley, M. M. & Lang, P. J. Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings. 1–45 (1999).

26. Radloff, L. S. The CES-D scale: A self-report depression scale for research in the general population. *Appl. Psychol. Meas.* **1,** 385–401 (1977).

27. Lewinsohn, P. M., Seeley, J. R., Roberts, R. E. & Allen, N. B. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychol. Aging* **12,** 277–87 (1997).

28. Unsworth, N., Heitz, R. P., Schrock, J. C. & Engle, R. W. An automated version of the operation span task. *Behav. Res. Methods* **37,** 498–505 (2005).

29. Dobson, D. Snood. *Snood LLC* (2011). Available at: snoodworld.com.

30. Smith, N. K. *et al.* Being bad isn't always good: affective context moderates the attention bias toward negative information. *J. Pers. Soc. Psychol.* **90,** 210–20 (2006).

31. Curran, T. Effects of attention and confidence on the hypothesized ERP correlates of recollection and familiarity. *Neuropsychologia* **42,** 1088–1106 (2004).

32. Woodruff, C. C., Hayama, H. R. & Rugg, M. D. Electrophysiological dissociation of the neural correlates of recollection and familiarity of recollection and familiarity. *Brain Res.* **1100,** 125–135 (2006).

33. Craik, F. I. M. & Tulving, E. Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* **104,** 268–294 (1975).

34. Smith, N. K., Cacioppo, J. T., Larsen, J. T. & Chartrand, T. L. May I have your attention, please: Electrocortical responses to positive and negative stimuli. *Neuropsychologia* **41,** 171–183 (2003).

35. Yonelinas, A. P. The nature of recollection and familiarity: A review of 30 years of research. *J Mem Lang* **46,** 441–517 (2002).

36. Wolk, D. A. *et al.* ERP correlates of recognition memory: Effects of retention interval and false alarms. *Brain Res.* 148–162 (2006). doi:10.1016/j.brainres.2006.04.050

37. Yu, S. S. *et al.* Electrophysiological dissociation of the neural correlates of

recollection and familiarity. *Brain Res.* **28,** 74–84 (2000).

38. Kaida, K. *et al.* Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clin. Neurophysiol.* **117,** 1574–1581 (2006).

39. Trujillo, L. T., Allen, J. J. B., Schnyer, D. M. & Peterson, M. a. Neurophysiological evidence for the influence of past experience on figure – ground perception. *J. Vis.* **10,** 1–21 (2010).

40. Nichols, T. E. & Holmes, A. P. Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Hum. Brain Mapp.* **15,** 1–25 (2002).